Measuring Shape-Dependent Looping Probability of DNA

Tung T. Le and Harold D. Kim*
School of Physics, Georgia Institute of Technology, Atlanta, Georgia

ABSTRACT Recently, several studies have shown that short doubled-stranded DNA (dsDNA) loops more readily than the wormlike chain model predicts. In most of these experiments, the intrinsic bendedness of dsDNA, which in theory can dramatically influence looping dynamics, was either avoided or unaccounted for. To investigate the effect of the shape of dsDNA on looping dynamics, we characterized the shapes of several synthetic dsDNA molecules of equal length but different sequences using gel electrophoresis. We then measured their looping rates using a FRET (Förster resonance energy transfer)-based assay and extracted the looping probability density known as the J factor (j_M). We also used, for comparison, several dinucleotide angular parameter sets derived from the observed electrophoretic mobility to compute the j_M predicted by the wormlike chain model. Although we found a strong correlation between curvature and j_M, the measured j_M was higher than most dinucleotide model predictions. This result suggests that it is difficult to reconcile the looping probability with the observed gel mobility within the wormlike chain model and underscores the importance of determining the intrinsic shape of dsDNA for proper theoretical analysis.

INTRODUCTION

DNA is arguably one of the most important molecules in science (1). Besides its biological importance as a carrier of the genetic code, it has become a versatile molecular building block for self-assembly of submicron structures. Therefore, understanding the mechanical properties of DNA is of fundamental importance in basic sciences and engineering applications. Although the typical double-stranded DNA (dsDNA) structure is portrayed as a double helix, dsDNA can be intrinsically curved, and can bend, twist, stretch, and/or melt due to thermal fluctuations. Despite this complexity, mechanical properties of DNA at large length scales can be well described by the wormlike chain model. According to this model, thermal energy on average bends dsDNA through an angle of one radian over ~50 nm (~147 bp), which is known as the persistence length. Therefore, loop formation of a 147-bp-long dsDNA which requires 2π bending over one persistence length would be rare.

However, looping of dsDNA on this length scale often plays an important role in biology (2), for example, in packaged DNA of bacterial viruses (3), transcription-factor mediated DNA looping (4), and winding of DNA in nucleosomes (5). Understanding dsDNA looping at small length scales demands a more detailed and thorough examination of local structural aberrations, which would be averaged out over large length scales. These aberrations can stem from static variations in mechanical properties such as stiffness and curvature, or from dynamic defects in the double helical structure, such as kinks and bubbles. There may also exist unknown mechanisms that cause elastic behavior of dsDNA at short length scales to deviate from the large-scale average behavior.

Spontaneous looping dynamics of dsDNA has been traditionally measured by ligase-dependent cyclization (6). In this assay, dsDNA molecules with sticky (cohesive) ends are circularized or dimerized by DNA ligase. By comparing the rates of circle and dimer formation, the looping efficiency can be quantified in terms of an effective molar concentration of one end of the DNA in the vicinity of the other end, which is known as the J factor (j_M). Widom et al. showed that the j_M of dsDNA shorter than the persistence length was much higher than predicted by the wormlike chain model (7). Although this conclusion was questionable due to the high ligase concentration used (8), several other groups also reported that dsDNA loops more frequently than a wormlike chain when different methods are employed (9–13). However, most of these kinetics studies involved proteins that could affect mechanical properties of dsDNA due to their nonspecific binding (14,15). To overcome this issue, Vafabakhsh and Ha used a FRET-based single-molecule assay to measure looping dynamics of short dsDNA in the absence of proteins (16). They measured j_M to be a few orders of magnitude higher than the wormlike chain model prediction for $<$100-bp dsDNAs.

In most dsDNA looping studies, the j_M-versus-length relationship was used as the litmus test for the wormlike chain model. The intrinsic bendedness of dsDNA was usually avoided or ignored, because permanent bendedness of dsDNA can in principle lead to abnormally high j_M at short length scales (17). Moreover, it is difficult to experimentally determine the exact shape of dsDNA in the ground state to use in j_M calculation. This poses the question: to what extent does curvature affect the looping dynamics of dsDNA? Here we designed several dsDNAs of equal length but different shape and used a similar FRET-based assay to Vafabakhsh and Ha (16) to measure their j_M. To compute j_M predicted by the wormlike chain model, we used the dinucleotide

Submitted January 26, 2013, and accepted for publication March 18, 2013.
*Correspondence: harold.kim@physics.gatech.edu
Editor: Taekjip Ha.
© 2013 by the Biophysical Society
0006-3495/13/05/2068/9 $2.00
http://dx.doi.org/10.1016/j.bpj.2013.03.029

chain model with angular parameters optimized by the gel electrophoretic mobility of these dsDNAs. We showed that measured j_M values were higher than in silico predictions in most cases. Our results suggest that mechanics of large-angle dsDNA bending cannot be reconciled with that of small-angle dsDNA bending at least in the dinucleotide chain framework.

MATERIALS AND METHODS

Experimental Design

We used two kinds of sequences: repeating artificial sequences and nonrepeating genomic sequences. The repeating sequences were variations of sequences chosen from a large-scale nucleosome occupancy study by Kaplan et al. (18). All sequences, both artificial and genomic, were terminated at each end with a common 18-bp-long adaptor sequence for subsequent polymerase chain reactions (PCRs). The repeating sequences were commercially synthesized by GenScript (sequences are given in the Supporting Material) and amplified using the following sets of primers synthesized by IDT:

F1: /5Cy3/**GGTAAATTCA**C**C**AACAACGAGCTCGAATG,
F1′: /5Cy3/**TAAATTC**CTACAACAACGAGCTCGAATG,
R1: GTGCCAGCAACAGATAGC,
F2: CAACAACGAGCTCGAATG,
R2: /5BioTEG/GAAACATA**G**/iCy5/**GAATTTACC**GTGCCAGCAAC AGATAGC.

For time efficiency, we used the backbone-incorporated Cy5 as the FRET acceptor instead of internally labeling a modified base with Cy5 in a separate step. We performed two separate PCRs on each sequence, one with primers F1 and R1 to integrate Cy3, and the other with primers F2 and R2 to integrate Cy5 and biotin. The 10-base-long complementary regions in F1 and R2 stabilize DNA in the looped state (shown in bold). The underlined A in primer F1 is a spacer that compensates for the internally linked Cy5 (iCy5) on R2. We also used F1′ with a shorter complementary region to increase the number of looping events. To construct the dsDNA carrying the FRET pair and biotin, we mixed the two PCR products with an excess amount of Cy3-carrying strand (molar ratio 4:1) and allowed strand exchange by heating and cooling the sample (incubation at 98.5°C for 1.5 min, gradual cooling to 5°C with ramp rate of 0.1°C/s, and incubation at 5°C for 2 h). More than 90% of surface-immobilized molecules carried Cy5, as expected, and $\geq$50% of them carried a Cy3 partner as a result of efficient strand exchange.

DNA electrophoresis

To measure curvature of dsDNA, we implemented polyacrylamide gel electrophoresis, similar to previously published protocols (29.2:0.8 acryl/bis, 5% in TBE buffer, PH 8.4, run at 5–8 V/cm in 4°C) (19,20). We compared the mobility of a DNA against the bands of a 1 kb DNA ladder (Fermentas) and interpolated its apparent length. The apparent size of the DNA (R) was defined as the ratio of its apparent length to its real length.

Optimizing helical parameters

We used the Levenberg-Marquardt nonlinear-least-squares algorithm in MATLAB (The MathWorks, Natick, MA) to find helical parameters that best predict the mobility of the seven artificial dsDNAs. We optimized 22 variables in total: 20 angular parameters corresponding to roll and tilt angles for all 10 unique dinucleotides and two additional parameters (a and b) for the linear relationship between the apparent size (R) and the central dispersion (σ^2):$R = a\sigma^2 + b$. We used 0.000156 and 0.97 as the initial values of a and b, as in the publication by De Santis et al. (21), and initialized the angular parameters based on seven different helical parameter sets, termed BT (22), OZ (23), CD (24), AS (25), LB (26), CHARMM27 (27), and PARMBSC0 (27). The twist angle was relatively insensitive to this optimization and was thus kept fixed (21). The three-dimensional dsDNA structure of a sequence was determined by the 30 angular variables, from which the central dispersion was calculated as published (21).

Fluorescence microscopy and image acquisition

For single-molecule imaging, we used an objective-type total internal reflection setup on a commercial microscope (IX81, Olympus, Melville, NY), as described (28). For FRET measurement, Cy3 (donor) was excited with a 532-nm laser (NT66-968, B&W Tek, Newark, DE). Cy5 (acceptor), when mentioned, was directly excited with a 640-nm laser (CUBE 640-30FP, Coherent, Santa Clara, CA). The donor and acceptor emissions were split into two separate paths by a dichroic mirror (FF545/650-Di01-25×36, Semrock, Rochester, NY). After passing through long-pass filters (LM01-552 for donor emission, FF650-Di01 for acceptor emission; Semrock), they were merged by the second dichroic mirror and projected on two halves of an EMCCD (DU-897ECS0-# BV, Andor). The laser power for Cy3 excitation was ~60 μW coming out of the objective. Binned images (2×2) were acquired at 25 frames/s with 40-ms exposure using our in-house software.

Glass slides and coverslips were cleaned according to a protocol of 20-min sonication in 10% Alconox, 10-min sonication in purified water, 15-min sonication in acetone, and 20-min sonication in 5 M KOH. Cleaned slides and coverslips were stored in purified water for long-term use. For single-molecule immobilization, both slides and coverslips were treated with polyethylene glycol polymer mixture (1:80 w/w ratio of biotin-(polyethylene glycol) PEG-silane, molecular weight 3400/mPEG-silane, molecular weight 2000 (w/w), Laysan Bio, Arab, AL) in 0.1 M sodium bicarbonate solution. After being rinsed with distilled water, they were dried in open air and assembled into a flow cell with multiple channels formed by double-stick tape. Immediately before each experiment, we injected 10 μl of Neutravidin solution (100 μg/ml) into the channel and waited for two minutes before washing with T50 buffer (10 mM Tris-HCl, 50 mM NaCl, pH 7.0). DNA molecules were introduced at 50–100 pM for specific binding to the PEG-coated surface through Neutravidin-biotin interaction. Based on our design, the presence of Cy3 on the surface indicates the presence of Cy5 in the same molecule. The channel was finally flushed with the imaging buffer, which contained an oxygen scavenging system (100 mM PCD, 5 mM PCA, and 1 mM Trolox, slightly modified from Aitken et al. (29)) and varying concentrations of NaCl (100–500 mM). All measurements took place at room temperature (22°C).

Data analysis

The movies, captured as a series of 256 × 256 images, were processed by a MATLAB code to generate single-molecule FRET traces and dwell-time histograms. The FRET traces were filtered using a two-point sliding average. A typical FRET histogram showed two peaks that corresponded to the looped state (high FRET value) and the unlooped state (low FRET value). Each histogram was fitted using a double Gaussian function, and the intersection between the two Gaussian curves was used as a hard threshold between the two states. Dwell times in each state were collected from traces showing switching behaviors. For each DNA sequence, ~500 traces from multiple viewing fields were used to generate dwell-time histograms.

In this study, we either reported the lifetimes of the unlooped and the looped states, which are the normal mean dwell times measured in most single-molecule studies, or the mean first looping time. The first looping

time is defined as the first passage time for an unlooped molecule to loop for the first time based on the observed FRET signal. Thus, it can be measured from molecules that exist in the unlooped state at the beginning of observation. For this measurement, we used molecules with a shorter, 7-bp complementary region so that the majority of the molecules were prepared in the unlooped state (~99%). Molecules that were either missing one of the two dyes or in the looped state (high FRET) at the beginning were discarded from analysis. Molecules that underwent Cy3 or Cy5 photobleaching before the first looping event were also discarded from analysis. Since the high-FRET state with the 7-bp complementary region was very short-lived (<1 s), Cy5 photobleaching was extremely rare. The typical photobleaching lifetime of Cy3 in these experiments was ~30–40 min, which was substantially longer than all looping times measured in our study.

To obtain the lifetime of the unlooped state or the mean first looping time, we plotted the number of molecules with both Cy3 and Cy5 that survived in the unlooped state as a function of time. We then fitted this decay curve using a double-exponential function with a base line: $N(t) = N_\infty + N_f \exp(-k_f t) + N_s \exp(-k_s t)$. Although the reason for this biphasic decay is not known, we suspect that molecules with small curvature can bend either inward or outward to form a loop, which might lead to double-exponential looping kinetics. In this equation, N_∞ accounts for the number of dysfunctional dsDNA molecules that are not able to loop for unknown reasons. The mean first looping time was calculated as $\langle t \rangle = N_f/(k_f(N_f+N_s)) + N_s/(k_s(N_f+N_s))$. The only difference between the two was whether the dwells in the unlooped state were synchronized on their unlooping transitions or not. The looping rate is defined as the inverse of the mean first looping time.

Numerical prediction of the J factor

We randomly sampled numbers from a Gaussian distribution, multiplied them with the roll, tilt, and twist fluctuation parameters, and added them to the ground-state roll, tilt, and twist angles to obtain roll, tilt, and twist angles for each dinucleotide in thermal equilibrium. Using these angles, we generated coordinates of dsDNA configurations (30). We used the free boundary condition where all possible angles between the two ends are allowed. The histogram of end-to-end distances (r) was normalized and divided by $4\pi r^2$ to obtain $P_{eq}(r)$, the probability density of the end-to-end distance being equal to r. j_M was obtained by estimating $P_{eq}(r)$ at 5.5 nm and converting it to molar concentration. Typically, 10^5 configurations were enough to yield reliable extrapolation of j_M for >150-bp dsDNA, and 10^6 configurations for <150-bp dsDNA.

RESULTS

FRET fluctuations reflect looping dynamics of dsDNA

To detect dsDNA looping, we designed DNA molecules with two regions: the looping region and the capturing region. The looping region is a 186-bp-long duplex (unless stated otherwise), and the capturing regions are 7- or 10-base-long single-stranded overhangs at both ends that can anneal to each other upon looping. One of the capturing regions also contains a biotin for surface immobilization (Fig. 1 *A*). The length of the capturing regions was chosen so that the looped state was stable enough to be detected by FRET, and unstable enough for the DNA to undergo reversible looping and unlooping.

We observed anticorrelated fluctuations of Cy3 (donor) and Cy5 (acceptor) from surface-immobilized dsDNA molecules (Fig. 1 *B*) due to reversible looping and unlooping of the DNA. We rule out blinking of the acceptor or switching of the acceptor dipole orientation as the cause of the fluctuations, because 1), Cy5 was fluorescently active in the low-FRET state, as demonstrated by the unintentional weak excitation of Cy5 by the 532-nm laser; 2), the observed intensity changes of the dyes in the capturing regions depended on the DNA sequence of the looping region; and 3), the mean dwell time (lifetime) of the looped state correlated with the number of complementary bases in the capturing region (Fig. S1).

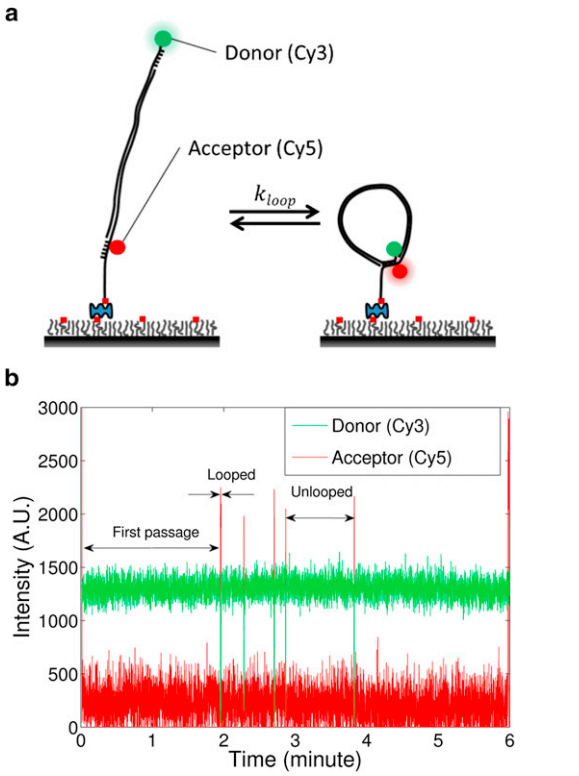

FIGURE 1 (*A*) Single-molecule FRET experiment. dsDNA molecules labeled with Cy3 (*green*) on one side and Cy5 (*red*) on the other are immobilized on the PEG-coated surface. Reversible looping and unlooping of dsDNAs result in high-FRET (looped) and low-FRET (unlooped) states. (*B*) A typical single-molecule time trace of donor (*green*) and acceptor (*red*) fluorescence intensities. The high- and low-FRET states are identified as the looped and unlooped states. Also shown are the first looping time and dwell times in the looped and unlooped states. See main text for definitions of these terms. At the beginning and end of the acquisition, 640-nm laser was switched on briefly to examine the presence of Cy5 in each molecule.

The lifetimes of the looped and unlooped states also changed with salt concentration (Fig. S2). The looped-state lifetime increased with salt, presumably because salt stabilizes base pairing. The unlooped-state lifetime decreased with increasing salt concentration, consistent with salt promoting annealing between the capturing overhangs. Salt can also accelerate looping by increasing dsDNA flexibility, but this effect is likely small, because the negatively charged

phosphate groups are almost completely screened in the range of concentrations tested (145–505 mM NaCl) (31,32). We used 505 mM NaCl as the standard concentration for all other experiments of this study.

We decreased the length of a dsDNA whose sequence was derived from the yeast genome from 186 to 102 bp and observed a 20-fold increase in the unlooped-state lifetime (Fig. 2 *A*). We emphasize that even if we had assumed that all molecules could loop eventually by fitting the decay curves with a zero-baseline exponential function, the change in the looping rate between 102 bp and 186 bp would still be <20-fold. To confirm that the measured looping kinetics was not affected by the acquisition time, we increased the acquisition time from 15 min to 45 min and measured the unlooped-state lifetime of the 102bp dsDNA. As is evident from Fig. S3, the decay curves are nearly indistinguishable.

In a ligation-based DNA cyclization assay that covalently links the 5′-end and the 3′-end of a dsDNA, j_M depends on the probability that the helical phases of the two ends match each other. This dependence results in oscillation of j_M as a function of length, whose peak-to-peak change can be as much as 100-fold at 180 bp according to the twisted wormlike chain model (33). To investigate whether our measured looping rate showed similar helical phase dependence, we measured looping rates of dsDNAs varying in length from 176 bp to 186 bp. We found that the looping rate changed no more than 1.3 times over this one helical period (Fig. 2 *B*). This result suggests that annealing between the two overhangs might not require as strict a helical phase match as ligation does in a cyclization assay.

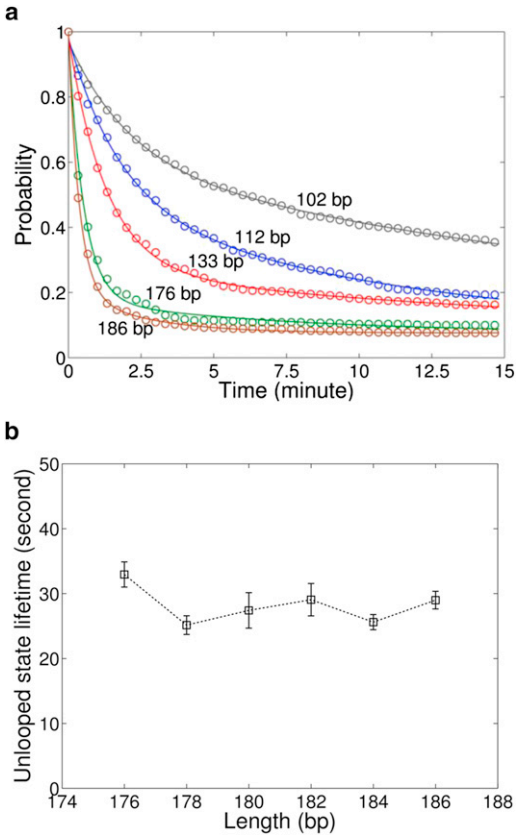

FIGURE 2 (*A*) Decay curves of the unlooped-state population as a function of length (*gray*, 102 bp; *blue*, 112 bp, *red*, 133 bp; *green*, 176 bp; *brown*, 186 bp). The decay curves were fitted using a double-exponential function with a fixed baseline extracted from the 186-bp data. (*B*) The unlooped-state lifetime of the genomic sequence DNA measured over one helical period. The small variation in the lifetime implies weak torsional stress for the dsDNA in the looped state.

Conversion of looping rate to j_M

The relationship between the equilibrium probability distribution of chain conformations and looping kinetics was given by Szabo et al. (34). According to this relationship, the mean first passage time between the two reactive terminal sites of a polymer chain is given by

$$\tau = \int_a^L dx\left[D(x)P_{eq}(x)\right]^{-1}\left(\int_x^L dy P_{eq}(y)\right)^2 + \frac{1}{\kappa_a P_{eq}(a)}, \quad (1)$$

where $P_{eq}(r)$ is the equilibrium probability density of the end-to-end distance being equal to r, L is the contour length of the polymer, D is the diffusion coefficient, and a is the inner boundary radius. κ_a reflects the trapping reaction between the two ends when $r = a$. In the case of a perfectly absorbing boundary condition ($\kappa_a \to \infty$), the second term vanishes.

We estimated the upper limit of this mean first passage time using an analytical expression of $P_{eq}(r)$ derived for a wormlike chain (35). The diffusion coefficient is estimated to be ~2.0×10^8 nm^2/s by treating the overhangs as 10-base-long rigid rods in water (36). The boundary radius is assumed to be ~5 nm. Using these values, the first term of Eq. 1 is ~5.5 ms, which is 10^2- to 10^3-fold shorter than our measured looping time. Other studies using molecular dynamics simulations or an analytical calculation also yielded a much shorter first passage time with efficient trapping (37–39). Therefore, the measured looping time likely reflects the second term in Eq. 1, in which case the inverse of the mean first passage time is directly proportional to $P_{eq}(r)$. Moreover, if the measured looping time reflects the first term in the equation, we expect it to increase linearly with viscosity. However, the looping time varied only 1.2-fold between 4% and 30% glycerol, which indicates that the apparent looping is not diffusion-limited (data not shown).

j_M is thus related to the mean first looping time (τ_{loop}) by

$$j_M(a) = \frac{P_{eq}(a)}{4\pi a^2} = \frac{1}{\kappa_a(4\pi a^2)\tau_{loop}}. \quad (2)$$

To extract j_M, κ_a must be known. This parameter can be measured from free diffusion of one sticky end with respect

to the other in a volume V, where $P_{eq}(a) = 4\pi a^2/V$. Substituting this in Eq. 1 and assuming that annealing is not diffusion-limited, we obtain the annealing time (τ_{anneal})

$$\tau_{\text{anneal}} \approx \frac{1}{N\kappa_a\left(\frac{4\pi a^2}{V}\right)}. \quad (3)$$

Here, we used the fact that in the presence of N sticky ends, the annealing rate increases N-fold. Substituting κ_a from Eq. 3 into Eq. 2, we obtain

$$j_M(a) = \frac{N}{V}\frac{\tau_{\text{anneal}}}{\tau_{\text{loop}}} = c\frac{\tau_{\text{anneal}}}{\tau_{\text{loop}}}. \quad (4)$$

Therefore, j_M can be determined from three experimental measurables: concentration of sticky ends (c), the looping time, and the annealing time.

We realized this experiment by immobilizing the Cy5-containing single strand to the surface and introducing the Cy3-containing single strand at 50 nM (Fig. S4 *A*). Annealing of the Cy5 strand to the Cy3 strand resulted in detectable FRET bursts, and the first-order rate constant was measured to be $0.45 \pm 0.04 \times 10^6\ M^{-1}\ s^{-1}$ in 505 mM NaCl (Fig. S4 *B*). This value is comparable to the values reported by the Ha group ($0.78 \pm 0.07 \times 10^6\ M^{-1}\ s^{-1}$ in 1 M NaCl and $0.26 \pm 0.04 \times 10^6\ M^{-1}\ s^{-1}$ in 10 mM Mg^{2+}) that were obtained from dimerization kinetics between dsDNAs (16).

More curved molecules loop faster

Although measuring the length dependence of j_M has been used as the litmus test for the wormlike behavior, the interpretation of the result is nontrivial for a couple of reasons. First, changing the length changes not only the bending energy, but also the twisting energy, required for looping. Second, curvature can be introduced locally when basepairs are added. Therefore, we kept the length of dsDNAs the same while varying the curvature in a global, predictable manner. To build curved DNA molecules, we concatenated a 10-mer sequence multiple times, with a random nucleotide inserted every 20 bases. As a result, the 10-mer sequence repeats itself every helical repeat (10.5 bp), and the static curvature, or bendedness, of the 10-mer, if there is any, will add up constructively. We designed four such repeating sequences, termed sequences 1–4. Sequences 1–3 are more similar to each other than they are to sequence 4. Sequence 2 and 4 were chosen from a previous study by Kaplan et al. (18) and are supposed to possess opposite nucleosome affinity. We also designed three extra sequences with locally perturbed curvature: 1), we changed G to T in the two central 10-mers of sequence 3 (sequence 3-1); 2), we changed G to T in the terminal 10-mers of sequence 3 (sequence 3-2); and 3), we truncated the central 10-mer of sequence 2 by half to make an S-shaped molecule with an inflection point (sequence 2-1).

We checked overall curvature of these dsDNAs by polyacrylamide gel electrophoresis. The order of apparent size (R) for sequences 1–4 was $1 > 2 > 4 > 3$ (Fig. 3 *A*). Among the sequence 3 variants, the apparent size was 3-2 > 3-1 > 3. The S-shaped molecule exhibited an apparent size between 2 and 4. Empirically, a curved dsDNA migrates more slowly than a straight dsDNA (19,25). Therefore, we could conclude that sequence 1 is the most curved, and that sequence 3 is the straightest. We measured looping kinetics of these dsDNAs, and found an anticorrelation between the apparent size and the unlooped-state lifetime (Fig. 3 *B*). The strong anticorrelation among the repeat sequences suggests that the deflection angle increases monotonically in the order $3 \rightarrow 4 \rightarrow 2 \rightarrow 1$ without exceeding 2π. Despite locally perturbed curvature, other sequences also followed the anticorrelation between apparent size and looping time, although the S-shaped molecule (sequence 2-1) deviated notably from the overall trend.

Estimating helical parameters from gel mobility

A curved dsDNA costs less bending energy to loop than a straight one and thus has a higher j_M. To compare j_M between different dsDNAs, one must know their intrinsic shape. Koo and Crothers obtained a relationship between the deflection angle of a dsDNA and its apparent size on a gel (19). However, the absolute deflection angle of a reference dsDNA was determined from its j_M according to the wormlike chain model. Since we wanted to test the wormlike

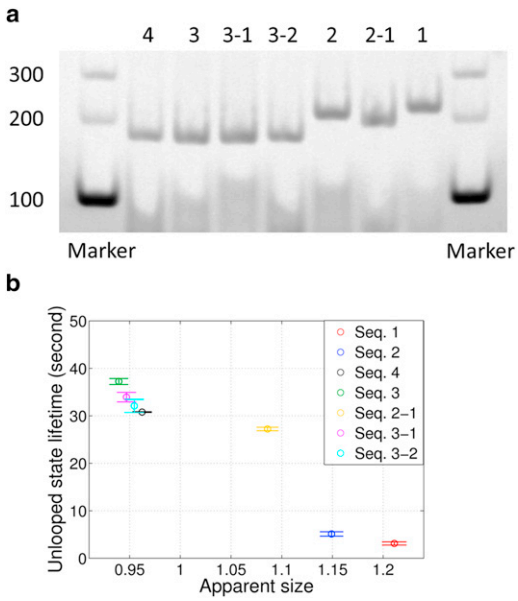

FIGURE 3 (*A*) Polyacrylamide gel electrophoresis pattern of synthetic DNAs. From left to right, the lanes contain 1 kb marker, sequence 4, sequence 3, sequence 3-1, sequence 3-2, sequence 2, sequence 2-1, sequence 1, and 1 kb marker. The exact sequences of these dsDNAs are given in the Supporting Material. (*B*) The relationship between the unlooped-state lifetime and the apparent size.

chain model, we used the relationship between the apparent size and the small-angle fluctuation of the DNA helical axis known as the central dispersion (σ^2) (25). According to Scipioni and co-workers, the apparent size of a dsDNA is related to the Boltzmann probability that its curvature is straightened out in thermal equilibrium and is thus proportional to its central dispersion.

In the dinucleotide model, there are 30 rotational parameters that define the relative geometry of 10 unique dinucleotides. Thus, finding these parameters that satisfy the linear relationship between apparent sizes of the seven dsDNAs and their central dispersion values is an underdetermined problem. Therefore, using nonlinear least-squares fitting, we optimized seven different helical parameter sets determined from crystal structures, gel electrophoresis, or MD simulations (22–27). The salt concentration we used is higher than what is typically used for gel electrophoresis or MD simulation but lower than that for crystallization. Using these different parameter sets can thus account for potential variability in the intrinsic shape of dsDNA arising from different salt conditions. This method enables us to obtain the most compatible adjustments of these seven dinucleotide parameter sets from the literature to match the observed gel mobility pattern without a priori assumption about the shape of any DNA molecule. The shapes of dsDNAs predicted by two of the optimized parameter sets are presented in Fig. S5.

Comparison of experiment with wormlike chain simulation

To compute j_M, we used the optimized sets of angular parameters with equal bending rigidity parameters in both roll and tilt directions that correspond to bending and twisting persistence lengths of 49.7 nm and 65 nm, respectively (17). Because the bending energy of a dinucleotide is proportional to the square of its angular displacement, the angular displacement is normally distributed in thermal equilibrium. Hence, we randomly sampled angular displacement values from a Gaussian distribution and built 10^5-10^6 chain configurations for each dsDNA. The number of chains whose end-to-end distance is within the annealing range (5.5 nm) is very small, especially for straighter dsDNAs. For example, sequence 4 typically yields none or a few loops out of 2 million trials. Therefore, estimation of j_M at a small distance is error-prone. To remedy this problem, we extrapolated the dependence of the logarithm of j_M on the end-to-end distance to obtain j_M at small distances (40). We used a rational function to fit this curve: $y = ax + b + c/(x-d)^2$, where b corresponds to j_M near zero end-to-end distance. An example of this extrapolation is shown in Fig. S6.

As shown in Fig. 4, in most combinations of DNA sequence and helical parameter set, the predicted j_M was lower than its experimental counterpart. Although the parameter set derived from the study of Bolshoy and Trifonov (BT) (22) yields a higher j_M value than the experimental value for curved dsDNAs, it significantly overestimates the fold change in j_M. We note that this parameter set tends to predict larger deflection angles compared to others (Fig. S5). All parameter sets were able to predict the increasing order of j_M. The ones that predict larger deflection angles tend to predict a larger change in j_M between the most curved and the straightest dsDNAs than measured. The ones that predict a similar change in j_M, however, underestimate j_M overall. This analysis suggests that it is difficult to reconcile the looping probability with the observed gel mobility within the wormlike model.

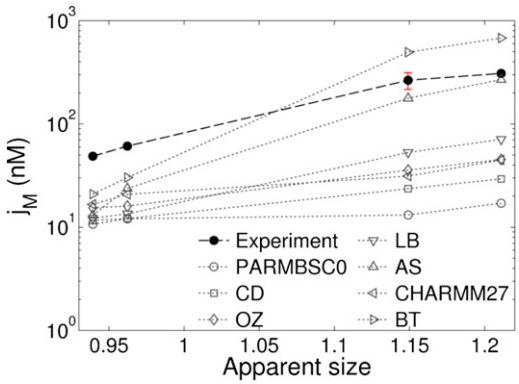

FIGURE 4 Dependence of j_M on curvature. j_M of dsDNAs uniformly curved by design (sequences 1–4) were experimentally determined according to Eq. 4. The mean values are shown as solid circles, with the error bars (representing the mean ± SE (n = 3)) in red. For comparison, j_M values calculated from simulation using seven different parameter sets are shown on the same plot. Except for the two values obtained using the BT parameter set (right-pointing triangles), all j_M values are predicted to be lower than measured according to the dinucleotide chain model.

Looping probability of a genomic sequence DNA as a function of length

Using the same parameter sets, we also computed the j_M of the genomic-sequence dsDNA as a function of length. As shown in Fig. 5, the BT parameter set produces j_M that closely matches the experimental data. CHARMM27 also works well, except at 112 bp. Other parameter sets can produce the overall rate of increase in j_M but underestimate the absolute value of j_M. We note that BT and CHARMM27 parameter sets can well predict the measured j_M values, because they also predict relatively large curvatures for the genomic sequence dsDNAs used (Fig. S7). Moreover, even for the CD and LB parameter sets, whose predictions deviate the most from experimental values, the discrepancy is not more than 10-fold between 100 and 200 bp. In comparison, the j_M of an intrinsically straight, isotropic chain, calculated from an analytical estimation by Douarche and Cocco (41) (Fig. 5, *green curve*), is substantially lower than the measured j_M, and the gap between them widens

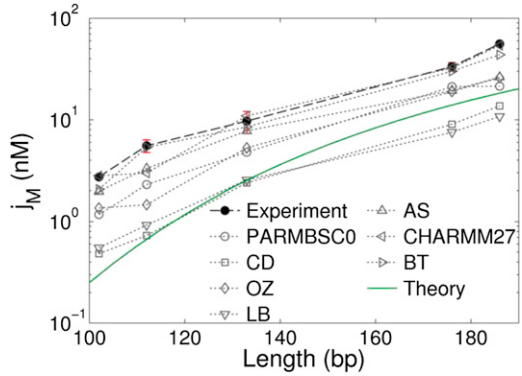

FIGURE 5 Dependence of j_M on length. j_M of the genomic sequence dsDNAs with varying length (sequences 102, 112, 133, 176, and 186) were experimentally determined according to Eq. 4. The mean values of three independent measurements are shown as solid spheres, with the error bars (representing the mean ± SE) shown in red. For comparison, j_M values calculated from simulation using seven different parameter sets are shown on the same plot. Most j_M values are predicted to be lower than measured according to the dinucleotide chain model. The green curve represents j_M of a uniform, isotropic wormlike chain calculated from an analytical estimation by Douarche and Cocco (42) with the constraint of a 5.5-nm end-to-end distance.

as length decreases. This result shows that the measured length dependence of j_M alone can be explained by the worm like chain model with a proper choice of dinucleotide parameters.

DISCUSSION

In summary, we studied looping kinetics of dsDNA of <200 bp using single-molecule FRET. The measured first looping times were much longer than the estimated first diffusive encounter time between the ends of the dsDNA, which allowed us to extract the equilibrium looping probability density, j_M, from the looping rate. j_M increased with curvature and decreased with length, as expected. We computed the j_M using helical parameter sets consistent with the observed gel mobility and found that it was lower than the measured j_M in most cases. We also demonstrate that it is difficult to test the wormlike chain model of dsDNA with j_M versus length measurement alone, as some parameter sets could reproduce the measured j_M between 100 and 200 bp. Hence, we propose the experimental comparison between gel mobility and looping probability of dsDNA at a fixed length as a powerful alternative approach to test consistency of the wormlike chain model.

In a ligase-based DNA cyclization assay, j_M is defined as the concentration of monomers that results in the same fraction of dimers and circles in equilibrium, and it can be estimated from the transient first-order rates of circle and dimer formation (42). However, this measured j_M is sensitive to the ligase concentration used and can be overestimated, as pointed out by Du et al. (8). Furthermore, nonspecific binding of DNA ligase to DNA can also affect j_M (14). In contrast, our FRET-based looping assay, similar to that of Vafabakhsh and Ha (16), is free from these concerns, and therefore can become a reliable alternative to measure dsDNA looping dynamics.

In this study, we used the mean first looping time (τ_{loop}) for theoretical comparison, for two reasons. First, Eq. 1 formulates looping kinetics in terms of the mean first passage time by averaging the first passage time over the unsynchronized initial end-to-end distribution of a polymer (34). Second, using the mean first looping time (τ_{loop}) instead of the unlooped-state lifetime reduces the risk of overestimating the frequency of rare events. For example, a routine dwell-time analysis includes events that are observed during the finite acquisition time only; it excludes slow events that could have been observed had the acquisition time been longer. Also, a long exposure to laser illumination required to capture rare events can elevate the sample temperature, thus accelerating the apparent looping rate.

The computation of j_M by simulation requires geometric constraints for the looped state. In a ligation-based DNA cyclization assay, j_M exhibits length-dependent oscillation, indicative of torsional stress involved in covalent loop closure. This helical phase dependence was also observed in other single-molecule dsDNA looping studies (16,43). However, we did not observe such oscillation in the looping rate over a 10-bp window. We note two main differences between our study and those studies (16,43). First, our dsDNAs are slightly longer, and therefore, the amplitude of j_M oscillation is expected to be smaller according to the twisted wormlike chain model. Second, our dsDNA has a gap of a few bases between opposite ends of one strand in the looped state, which prevents proper stacking of bases around it. In any case, because of the lack of helical phase dependence, we applied end proximity as the only constraint for j_M calculation.

In this study, dsDNA was attached to the surface through a terminal base. This raises the question of how the surface attachment would affect j_M. According to our simulation, the j_M of a dsDNA terminally anchored to the surface is lower than that of a dsDNA free in solution (to be addressed in a future publication). Therefore, the j_M values we report here should be taken as low estimates, implying that the discrepancy in j_M between a real dsDNA and a wormlike chain might be even larger. Another potential concern with this FRET-based assay is the presence of dye labels in the capturing region, which can interfere with cyclization. We stress that j_M was extracted by dividing the apparent cyclization rate by the rate of annealing between the two sticky ends, which was measured independently of looping. Therefore, the effect of dyes on cyclization, if any, is effectively normalized out by this procedure.

Our experimental finding that intrinsic curvature of dsDNA has a dominant effect on its looping probability seems to be at odds with a previous study by Vologodskaia and Vologodskii (44). There, it was reported that intrinsic

curvature has only a minor impact on cyclization. We point out that the lambda DNA used in that study does not have a repeating sequence feature that can give rise to a consistent direction of bendedness. Such randomized curvature does not lead to a dramatic increase in looping probability, as illustrated by the S-shaped dsDNA used in our study. Although the S-shaped dsDNA is highly curved, j_M is not relatively high, because the two halves of the molecule are curved in opposite directions.

Our results demonstrate that it is difficult to reconcile the energetics of small-angle dsDNA bending with that of large-angle or sharp bending. The angular parameters derived based on gel mobility of mildly curved dsDNAs are optimal for describing small-angle bending, and we show that they underestimate j_M that describes sharp bending. In gel electrophoresis, the gel network can be thought of as a free-energy barrier that limits movement of curved dsDNA. In this simple thermodynamic model, gel mobility can be theoretically interpreted as the probability of a curved dsDNA to be straightened (25). If the curvature of a dsDNA is relatively small, its mobility can be shown to be linearly proportional to the central dispersion of the dsDNA. The dsDNA molecules we used in this study show apparent sizes well within the range where the linear relationship is experimentally valid (25). If the curvature is too large, the molecule can entangle with the gel network in a complicated fashion and fail to obey the linear relationship. In this case, inferring the intrinsic shape of dsDNA from the absolute electrophoretic mobility may require more detailed theoretical models (45–47).

Our result is in qualitative agreement with a subelastic chain model where the bending energy of dsDNA has a less than quadratic dependence on the bending angle (48). The most discussed mechanism for subelastic dsDNA is the formation of defects such as kinks and bubbles in the dsDNA structure (49). However, the existence of stable kinks or bubbles in dsDNAs as long as 190 bp is very unlikely at the high salt concentrations we used (8). Another mechanism that does not require bubble formation comes from a recent molecular dynamics study (50). In that study, the authors proposed that counterions preferentially congregate on the concave side of the sharply bent dsDNA, and therefore, the free-energy cost of bending dsDNA increases linearly with bending angle at large bending angles. It is also possible that the dinucleotide chain model fails at large bending angles because nearest-neighbor effects become significant (51). Further studies on the effect of salt, temperature, and sequence on dsDNA looping will help elucidate the mechanism of extreme dsDNA flexibility.

SUPPORTING MATERIAL

Seven figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)00368-8.

We are especially grateful to Dr. Taekjip Ha for sharing his unpublished manuscript. We thank Dr. Stephen Harvey, Dr. Yongli Zhang, and Dr. Huan-Xiang Zhou for useful discussions. We thank James Waters for critical reading of the manuscript and providing preliminary data for the surface effect.

REFERENCES

1. Seeman, N. C. 2010. Nanomaterials based on DNA. *Annu. Rev. Biochem.* 79:65–87.
2. Garcia, H. G., P. Grayson, …, P. A. Wiggins. 2007. Biological consequences of tightly bent DNA: the other life of a macromolecular celebrity. *Biopolymers.* 85:115–130.
3. Baker, T. S., N. H. Olson, and S. D. Fuller. 1999. Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiol. Mol. Biol. Rev.* 63:862–922.
4. Bond, L. M., J. P. Peters, …, L. J. Maher, 3rd. 2010. Gene repression by minimal *lac* loops in vivo. *Nucleic Acids Res.* 38:8072–8082.
5. Andrews, A. J., and K. Luger. 2011. Nucleosome structure(s) and stability: variations on a theme. *Annu. Rev. Biophys.* 40:99–117.
6. Shore, D., J. Langowski, and R. L. Baldwin. 1981. DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl. Acad. Sci. USA.* 78:4833–4837.
7. Cloutier, T. E., and J. Widom. 2004. Spontaneous sharp bending of double-stranded DNA. *Mol. Cell.* 14:355–362.
8. Du, Q., C. Smith, …, A. Vologodskii. 2005. Cyclization of short DNA fragments and bending fluctuations of the double helix. *Proc. Natl. Acad. Sci. USA.* 102:5397–5402.
9. Cloutier, T. E., and J. Widom. 2005. DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proc. Natl. Acad. Sci. USA.* 102:3645–3650.
10. Wiggins, P. A., T. van der Heijden, …, P. C. Nelson. 2006. High flexibility of DNA on short length scales probed by atomic force microscopy. *Nat. Nanotechnol.* 1:137–141.
11. Yuan, C., E. Rhoades, …, L. A. Archer. 2006. Spontaneous sharp bending of DNA: role of melting bubbles. *Nucleic Acids Res.* 34:4554–4560.
12. Yuan, C., H. Chen, …, L. A. Archer. 2008. DNA bending stiffness on small length scales. *Phys. Rev. Lett.* 100:018102.
13. Han, L., H. G. Garcia, …, R. Phillips. 2009. Concentration and length dependence of DNA looping in transcriptional regulation. *PLoS ONE.* 4:e5621.
14. Yuan, C., X. W. Lou, …, L. A. Archer. 2007. T4 DNA ligase is more than an effective trap of cyclized dsDNA. *Nucleic Acids Res.* 35:5294–5302.
15. Manzo, C., C. Zurla, …, L. Finzi. 2012. The effect of nonspecific binding of lambda repressor on DNA looping dynamics. *Biophys. J.* 103:1753–1761.
16. Vafabakhsh, R., and T. Ha. 2012. Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization. *Science.* 337:1097–1101.
17. Czapla, L., D. Swigon, and W. K. Olson. 2006. Sequence-dependent effects in the cyclization of short DNA. *J. Chem. Theory Comput.* 2:685–695.
18. Kaplan, N., I. K. Moore, …, E. Segal. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature.* 458:362–366.
19. Koo, H. S., and D. M. Crothers. 1988. Calibration of DNA curvature and a unified description of sequence-directed bending. *Proc. Natl. Acad. Sci. USA.* 85:1763–1767.
20. Prosseda, G., A. Mazzola, …, B. Colonna. 2010. A temperature-induced narrow DNA curvature range sustains the maximum activity of a bacterial promoter in vitro. *Biochemistry.* 49:2778–2785.

21. De Santis, P., A. Palleschi, …, A. Scipioni. 1990. Validity of the nearest-neighbor approximation in the evaluation of the electrophoretic manifestations of DNA curvature. *Biochemistry.* 29:9269–9273.
22. Bolshoy, A., P. McNamara, …, E. N. Trifonov. 1991. Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. *Proc. Natl. Acad. Sci. USA.* 88:2312–2316.
23. Olson, W. K., A. A. Gorin, …, V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA.* 95:11163–11168.
24. Calladine, C. R., H. R. Drew, and M. J. McCall. 1988. The intrinsic curvature of DNA in solution. *J. Mol. Biol.* 201:127–137.
25. Anselmi, C., P. De Santis, …, A. Scipioni. 2002. From the sequence to the superstructural properties of DNAs. *Biophys. Chem.* 95:23–47.
26. Liu, Y., and D. L. Beveridge. 2001. A refined prediction method for gel retardation of DNA oligonucleotides from dinucleotide step parameters: reconciliation of DNA bending models with crystal structure data. *J. Biomol. Struct. Dyn.* 18:505–526.
27. Pérez, A., F. Lankas, …, M. Orozco. 2008. Towards a molecular dynamics consensus view of B-DNA flexibility. *Nucleic Acids Res.* 36:2379–2394.
28. Friedman, L. J., J. Chung, and J. Gelles. 2006. Viewing dynamic assembly of molecular complexes by multi-wavelength single-molecule fluorescence. *Biophys. J.* 91:1023–1031.
29. Aitken, C. E., R. A. Marshall, and J. D. Puglisi. 2008. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* 94:1826–1835.
30. Zhang, Y., and D. M. Crothers. 2003. Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.* 84:136–153.
31. Baumann, C. G., S. B. Smith, …, C. Bustamante. 1997. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. USA.* 94:6185–6190.
32. Guo, Z., C. H. Taubes, …, U. Mohanty. 2008. DNA on a tube: electrostatic contribution to stiffness. *J. Phys. Chem. B.* 112:16163–16169.
33. Shimada, J., and H. Yamakawa. 1984. Ring-closure probabilities for twisted wormlike chains. Application to DNA. *Macromolecules.* 17:689–698.
34. Szabo, A., K. Schulten, and Z. Schulten. 1980. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* 72:4350–4357.
35. Wilhelm, J., and E. Frey. 1996. Radial distribution function of semiflexible polymers. *Phys. Rev. Lett.* 77:2581–2584.
36. Nkodo, A. E., J. M. Garnier, …, G. W. Slater. 2001. Diffusion coefficient of DNA molecules during free solution electrophoresis. *Electrophoresis.* 22:2424–2432.
37. Podtelezhnikov, A. A., and A. V. Vologodskii. 2000. Dynamics of small loops in DNA molecules. *Macromolecules.* 33:2767–2771.
38. Jun, S., J. Bechhoefer, and B.-Y. Ha. 2003. Diffusion-limited loop formation of semiflexible polymers: Kramers theory and the intertwined time scales of chain relaxation and closing. *Europhys. Lett.* 64:420–426 (EPL).
39. Hyeon, C., and D. Thirumalai. 2006. Kinetics of interior loop formation in semiflexible chains. *J. Chem. Phys.* 124:104905, 104905–104914.
40. Hagerman, P. J. 1985. Analysis of the ring-closure probabilities of isotropic wormlike chains: application to duplex DNA. *Biopolymers.* 24:1881–1897.
41. Douarche, N., and S. Cocco. 2005. Protein-mediated DNA loops: effects of protein bridge size and kinks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 72:061902.
42. Taylor, W. H., and P. J. Hagerman. 1990. Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. II. NaCl-dependence of DNA flexibility and helical repeat. *J. Mol. Biol.* 212:363–376.
43. Laurens, N., D. A. Rusling, …, G. J. Wuite. 2012. DNA looping by FokI: the impact of twisting and bending rigidity on protein-induced looping dynamics. *Nucleic Acids Res.* 40:4988–4997.
44. Vologodskaia, M., and A. Vologodskii. 2002. Contribution of the intrinsic curvature to measured DNA persistence length. *J. Mol. Biol.* 317:205–213.
45. Haran, T. E., I. Cohen, …, U. Mohanty. 2003. Dynamics of curved DNA molecules: prediction and experiment. *J. Am. Chem. Soc.* 125:11160–11161.
46. Haran, T. E., I. Cohen, …, U. Mohanty. 2004. Characteristics of migration patterns of DNA oligomers in gels and the relationship to the question of intrinsic DNA bending. *J. Am. Chem. Soc.* 126:2372–2377.
47. Haran, T. E., and U. Mohanty. 2009. The unique structure of A-tracts and intrinsic DNA bending. *Q. Rev. Biophys.* 42:41–81.
48. Wiggins, P. A., and P. C. Nelson. 2006. Generalized theory of semiflexible polymers. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 73:031906.
49. Yan, J., and J. F. Marko. 2004. Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys. Rev. Lett.* 93:108108.
50. Spiriti, J., H. Kamberaj, …, A. van der Vaart. 2012. DNA bending through large angles is aided by ionic screening. *J. Chem. Theory Comput.* 8:2145–2156.
51. Lavery, R., K. Zakrzewska, …, J. Sponer. 2010. A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res.* 38:299–313.